

# SUBORDINATION OF THE RESOLVENT FOR A SUM OF RANDOM MATRICES

V. KARGIN

## Abstract

This paper is about the relation of random matrix theory and the subordination phenomenon in complex analysis. Previously, subordination was discovered in free probability theory by Voiculescu ([13]) and Biane ([1]). We find that approximate subordination holds for certain ensembles of random matrices. This allows us to prove an improved local limit law for eigenvalues and a delocalization result for eigenvectors of these ensembles.

## 1. INTRODUCTION

The operator-valued analytic function  $G_A(z) = (A - zI)^{-1}$  is called the resolvent of operator  $A$ . Much of the modern approach to random matrices is based on the analysis of how the resolvent of an  $N$ -by- $N$  random matrix  $A_N$  behaves when  $A_N$  is modified, for example when a column and a row is added. The results of this analysis are an important ingredient in recent investigations of eigenvalues of Wigner matrices ([4]). In this paper we investigate what happens with the resolvent of a matrix if an independent rotationally invariant random matrix is added to it. We find that the resolvent of the sum is (approximately) subordinated to the resolvent of the original matrix.

Subordination is an important concept in the theory of functions of complex variables. If  $f(z)$  and  $g(z)$  are two functions analytic in the upper half-plane  $\mathbb{C}^+ = \{z : \text{Im} z > 0\}$ , then  $f(z)$  is *subordinated* to  $g(z)$  if there exists an analytic function  $\omega(z) : \mathbb{C}^+ \rightarrow \mathbb{C}^+$ , such that  $f(z) = g(z + \omega(z))$  (see [9]). In this definition,  $f(z)$  and  $g(z)$  can be vector or even operator valued.

An indication of the relation between subordination and random matrices gleamed when Voiculescu discovered ([13]) that the Stieltjes transform of the free convolution of probability measures is subordinated to the Stieltjes transforms of each of the original measures. Since free convolution is closely related to sums of large random matrices, this indicated that subordination should hold for random matrices.

Voiculescu proved his subordination result under some non-degeneracy conditions, which were later removed by Biane in [1]. Moreover, Biane proved a significant strengthening of

---

*Date:* July 2011.

Department of Mathematics, Stanford University, Palo Alto, CA 94305, USA. e-mail: kargin@stanford.edu.

this result by extending it to resolvents of sums of free operators. Let us recall the form in which the subordination result was obtained by Biane.

Let  $\mathcal{A}$  be a von Neumann operator algebra with the normal faithful trace  $\tau$ , and let  $\mathcal{B}$  be a von Neumann subalgebra of  $\mathcal{A}$ . A conditional expectation  $\tau(\cdot|\mathcal{B})$  is a weakly continuous linear map  $\mathcal{A} \rightarrow \mathcal{B}$  determined by the following properties: (i)  $\tau(\mathbf{1}|\mathcal{B}) = \mathbf{1}$ , and (ii) if  $b_1, b_2 \in \mathcal{B}$ , then  $\tau(b_1 a b_2|\mathcal{B}) = b_1 \tau(a|\mathcal{B}) b_2$ .

Biane's result (cf. Theorem 3.1 in [1]) says that if two self-adjoint operators  $A, B \in \mathcal{A}$  are free, then the following identity holds for their resolvents:

$$\tau(G_{A+B}(z)|\mathcal{A}) = G_A(z + S(z)), \quad (1)$$

where  $\tau(\cdot|\mathcal{A})$  denote the conditional expectation on the subalgebra generated by operator  $A$ , and  $S(z)$  is a function analytic in the upper half-plane  $\mathbb{C}^+$ . This function  $S(z)$  depends only on  $A$  and  $B$ , maps  $\mathbb{C}^+$  to  $\overline{\mathbb{C}^+}$  and satisfies the relation  $S(i\eta)/i\eta \rightarrow 0$  as  $\eta \rightarrow \infty$ .

In other words,  $\tau(G_{A+B}(z)|\mathcal{A})$  is subordinated to  $G_A(z)$ .

Since the concept of freeness is closely related to asymptotic behavior of large independent random matrices, it is natural to ask whether subordination results hold in the context of random matrices.

More precisely, let  $A_N$  and  $B_N$  be two  $N$ -by- $N$  Hermitian matrices, and let  $H_N = A_N + U_N B_N U_N^*$ , where  $U_N$  is a random  $N$ -by- $N$  matrix with the Haar distribution on the unitary group. The resolvent of  $H_N$  is  $G_{H_N}(z) = (H_N - z)^{-1}$  and the resolvents of matrices  $A_N$  and  $B_N$  are defined similarly.

It is known that for large  $N$ , the behavior of operators  $A_N$  and  $U_N B_N U_N^*$  resembles the behavior of free operators. In particular, we can ask the question whether  $G_{H_N}(z)$  is subordinated to  $G_{A_N}(z)$  and  $G_{B_N}(z)$  for sufficiently large  $N$ .

Since matrices  $A_N$  and  $U_N B_N U_N^*$  provide only an approximation to free operators, a more cautious version of this question is whether subordination holds in an approximate sense. In other words, we are interested in the existence of an analytic function  $S_{B,N}(z) : \mathbb{C}^+ \rightarrow \Omega_N$  such that

$$\mathbb{E} G_{H_N}(z) = G_{A_N}(z + S_{B,N}(z)) + R_{A,N}(z), \quad (2)$$

where  $\Omega_N$  is a region that approaches  $\mathbb{C}^+$  as  $N \rightarrow \infty$ , and  $\|R_{A,N}(z)\| \rightarrow 0$  as  $N \rightarrow \infty$ . Here  $\mathbb{E}$  denote the expectation with respect to the Haar measure on unitary matrices  $U_N$ . (Of course, a similar question can be asked about subordination with respect to  $G_{B_N}$ . Since all considerations are the same, we will treat only the case of  $G_{A_N}$ .)

It turns out that a suitable candidate for such a function is the following expression:

$$S_{B,N}(z) := -\frac{\mathbb{E} f_{B_N}(z)}{\mathbb{E} m_{H_N}(z)},$$

where

$$\begin{aligned} m_{H_N}(z) &= N^{-1} \text{Tr} \left( \frac{1}{H_N - z} \right) \\ &= N^{-1} \text{Tr} (G_{H_N}(z)) \end{aligned}$$

is the *Stieltjes transform* of  $H_N$  and.

$$f_{B_N}(z) := N^{-1} \text{Tr} \left( B_N \frac{1}{H_N - z} \right).$$

( $S_{A,N}(z)$  and  $f_{A_N}(z)$  are defined similarly.)

For this choice of  $S_{B,N}(z)$ , it is known ([10]) that the norm of the error term  $\|R_{A,N}(z)\| \rightarrow 0$  as  $N \rightarrow \infty$  provided that  $\text{Im} z > \eta_0 > 0$ . Besides, it is clear that  $S_{B,N}(z)$  is analytic in  $\mathbb{C}^+$  since  $\mathbb{E}m_{H_N}(z)$  is the Stieltjes transform of a probability measure and therefore has no zeros in  $\mathbb{C}^+$ . Hence the main question is whether  $S_{B,N}(z)$  maps  $\mathbb{C}^+$  to  $\overline{\mathbb{C}^+}$ . Our first result is as follows.

**Theorem 1.** *Suppose  $A_N = B_N$  and  $\text{Im} z > 0$ . Then  $\text{Im} S_{B,N}(z) \geq 0$  and equality holds if and only if  $A_N$  and  $B_N$  are a multiple of identity matrix:  $A_N = B_N = cI_N$ .*

**Proof:** Definitions imply the following identity:

$$S_{A,N}(z) + S_{B,N}(z) = -z - \frac{1}{\mathbb{E}m_H(z)}.$$

By assumption,  $A_N = B_N$ , and therefore,

$$S_{A,N}(z) = S_{B,N}(z) = -\frac{1}{2} \left( z + \frac{1}{\mathbb{E}m_H(z)} \right).$$

Let  $\eta = \text{Im} z > 0$ . If  $\lambda \in \mathbb{R}$ , then  $(\lambda - i\eta)^{-1} \in \overline{D}_\eta$ , where  $\overline{D}_\eta$  is the closed disc with the diameter  $[0, i/\eta]$ . Since  $\mathbb{E}m_H(z)$  is a weighted average of  $(\lambda - i\eta)^{-1}$ , we conclude that  $\mathbb{E}m_H(z) \in \overline{D}_\eta$ . Therefore,  $\text{Im} (\mathbb{E}m_H(z))^{-1} \leq -\eta$  and  $\text{Im} \left[ z + (\mathbb{E}m_H(z))^{-1} \right] \leq 0$ . Therefore,  $\text{Im} (S_{A,N}(z)) = \text{Im} (S_{B,N}(z)) \geq 0$ .

The case of equality can occur only if  $\mathbb{E}m_H(z) \in \partial \overline{D}_\eta$  which implies that eigenvalues of  $m_H(z)$  are concentrated at a single point on the boundary of  $\overline{D}_\eta$  and this point does not depend on  $U_N$ . Therefore  $H_N = 2cI_N$  for a real  $c$ . This implies that  $A_N = B_N = cI_N$ .  $\square$

For the more general case of non-equal  $A_N$  and  $B_N$ , we will prove the following result.

**Theorem 2.** *Assume that  $\max \{\|A_N\|, \|B_N\|\} \leq K$  and that  $z = E + i\eta \in \Lambda_R := \{z : z > 0, |z| \leq R\}$ . Then for all  $N \geq c_1\eta^{-5}$ ,*

$$\min \{ \text{Im} (S_{A,N}(z)), \text{Im} (S_{B,N}(z)) \} \geq -\frac{c_2}{N\eta^7},$$

where  $c_1, c_2 > 0$  depend only on  $K$  and  $R$ .

Intuitively, this theorem says that for all sufficiently large  $N$ , the imaginary part of the subordination functions is almost non-negative. It is possible that in fact a stronger statement holds, namely that  $\min \{ \text{Im} (S_{A,N}(z)), \text{Im} (S_{B,N}(z)) \} \geq 0$  for all  $z \in \mathbb{C}^+$  and all sufficiently large  $N$  (which perhaps depend on  $\text{Im} z$ .) The author knows neither counterexamples to this stronger statement nor its proof. The proof of Thm. 2 will be given in the next section.

Theorems 1 and 2 allow us to improve the estimate of the error term in (2).

**Proposition 1.** *Let  $R_{A,N}(z)$  be defined as in (2). Assume that  $\max\{\|A_N\|, \|B_N\|\} \leq K$  and  $z = E + i\eta \in \Lambda_R = \{z : \eta > 0, |z| \leq R\}$ . Then,*

- (i) *if  $A = B$ , then  $\|R_{A,N}(z)\| \leq c/(N\eta^6)$ ,*
- (ii) *if  $A \neq B$ , then  $\|R_{A,N}(z)\| \leq c/(N\eta^6)$  provided that  $N\eta^8 \geq c_1$ ,*

*where the constants  $c$  and  $c_1$  depend only on  $K$  and  $R$ .*

Note that by using this result we can estimate  $R_{A,N}(z)$  at points arbitrarily close to the real axis provided that  $N$  is sufficiently large. This fact allows us to prove a local limit result for the convergence of eigenvalue counting measure.

Let us introduce additional definitions needed to formulate this result.

Let  $\lambda_1^{(H_N)} \geq \dots \geq \lambda_N^{(H_N)}$  denote eigenvalues of  $H_N$ , and define the *spectral measure* of  $H_N$  as  $\mu_{H_N} := N^{-1} \sum_{k=1}^N \delta_{\lambda_k^{(H_N)}}$ . Define  $\mu_{A_N}$  and  $\mu_{B_N}$  similarly to  $\mu_{H_N}$ , and assume that  $\mu_{A_N} \rightarrow \mu_\alpha$  and  $\mu_{B_N} \rightarrow \mu_\beta$  as  $N \rightarrow \infty$ , where  $\mu_\alpha$  and  $\mu_\beta$  are probability measures and convergence is the weak convergence of measures. Let  $\mathcal{N}_{\eta^*}(E)$  be the number of eigenvalues of  $H_N$  in the interval  $I^* = [E - \eta^*, E + \eta^*]$ . What can be said about  $\mathcal{N}_{\eta^*}(E) / (2\eta^* N)$  when  $N \rightarrow \infty$ ?

If  $\eta^*$  is fixed, then it is known ([12]) and ([11]) that the limit approaches  $\mu_\alpha \boxplus \mu_\beta(I^*)$ , where  $\boxplus$  denotes *free convolution*, a non-linear operation on probability measures introduced by Voiculescu in his studies of operator algebras. Local limit theorems address the question about what happens if  $\eta^*$  is not fixed but approaches 0 when  $N \rightarrow \infty$ .

We will formulate an answer in Theorem 3 below after giving some more definitions.

Let  $\mu_\alpha$  and  $\mu_\beta$  be two probability measures with bounded support, and let  $m_\alpha(z)$  and  $m_\beta(z)$  be their Stieltjes transforms. The system of equations:

$$\begin{aligned} m(z) &= m_\alpha(z + S_\beta(z)), \\ m(z) &= m_\beta(z + S_\alpha(z)), \\ -\left(z + \frac{1}{m(z)}\right) &= S_\alpha(z) + S_\beta(z). \end{aligned} \tag{3}$$

has a unique solution  $(m(z), S_\alpha(z), S_\beta(z))$  in the class of functions that are analytic in  $\mathbb{C}^+ = \{z : \text{Im} z > 0\}$  and have the following expansions at infinity:

$$\begin{aligned} m(z) &= -z^{-1} + O(z^{-2}), \\ S_{A,B}(z) &= O(1). \end{aligned} \tag{4}$$

The function  $m(z)$ , which we denote  $m_\boxplus$ , is the Stieltjes transform of a probability measure which is in fact the free convolution of measures  $\mu_\alpha$  and  $\mu_\beta$ . It is known ([1]) that  $\text{Im} S_\alpha(z)$  and  $\text{Im} S_\beta(z)$  are non-negative if  $\text{Im} z > 0$ .

**Definition 1.** *A pair of probability measures  $(\mu_\alpha, \mu_\beta)$  is called smooth at  $E$  if the limits of functions  $S_\alpha(E) := \lim_{\eta \downarrow 0} S_\alpha(E + i\eta)$  and  $S_\beta(E) := \lim_{\eta \downarrow 0} S_\beta(E + i\eta)$  exist and if both  $\text{Im} S_\alpha(E)$  and  $\text{Im} S_\beta(E)$  are strictly positive.*

**Definition 2.** A point  $E \in \mathbb{R}$  is called *generic with respect to measures*  $(\mu_\alpha, \mu_\beta)$  if the following equation holds:

$$k(E) := \frac{1}{m'_\alpha(E + S_\beta(E))} + \frac{1}{m'_\beta(E + S_\alpha(E))} - (E + S_\alpha(E) + S_\beta(E))^2 \neq 0. \quad (5)$$

We will say that  $(\mu_\alpha, \mu_\beta)$  is *well-behaved* at  $E$ , if  $(\mu_\alpha, \mu_\beta)$  is smooth at  $E$  and  $E$  is generic with respect to  $(\mu_\alpha, \mu_\beta)$ . It is known that if  $\mu_\alpha = \mu_\beta$  then  $(\mu_\alpha, \mu_\beta)$  is smooth at  $E$  if and only if the free convolution  $\mu_\alpha \boxplus \mu_\beta$  has a positive density at  $E$ . Hence, smoothness is a strengthening of the requirement that  $\mu_\alpha \boxplus \mu_\beta$  is absolutely continuous at  $E$ . The concept of genericity is needed to guarantee that the solution of the system (3) at  $E$  is stable with respect to small perturbation in the system. It comes from the requirement that the determinant of the Jacobian matrix of the system be different from 0 at  $E$ .

Next, let  $\mu_1$  and  $\mu_2$  be two probability measures on the real line, and let  $F_1(t)$  and  $F_2(t)$  be their cumulative distribution functions. The *Levy distance* between measures  $\mu_1$  and  $\mu_2$  is defined by the formula:

$$d_L(\mu_1, \mu_2) = \sup_x \inf \{s \geq 0 : F_2(x - s) - s \leq F_1(x) \leq F_2(x + s) + s\}.$$

The Levy distance is in fact a metric on the space of probability measures and the convergence with respect to this metric is equivalent to the weak convergence of measures.

Now we are ready to formulate the local limit theorem.

**Theorem 3.** Let  $1 \gg \eta^* \gg N^{-1/8}$ , and assume that (i)  $\max \{d_L(\mu_{A_N}, \mu_\alpha), d_L(\mu_{B_N}, \mu_\beta)\} \rightarrow 0$ , (ii) the pair of probability measures  $(\mu_\alpha, \mu_\beta)$  is well-behaved at  $E$ , (iii)  $\max \{\|A_N\|, \|B_N\|\} \leq K$  for all  $N$ . Let  $\rho_\boxplus$  denote the density of  $\mu_\alpha \boxplus \mu_\beta$ . Then,

$$\frac{N_{\eta^*}(E)}{2\eta^*N} \rightarrow \rho_\boxplus(E)$$

in probability.

Here notation  $\eta^* \gg N^{-1/8}$  means that  $\eta^*/N^{-1/8} \rightarrow \infty$  as  $N \rightarrow \infty$ .

The theorem improves the local limit law in [7], where it was found that it holds for the window size  $\eta^* \sim (\log N)^{-1/2}$ . The expected optimal result is  $\eta^* \sim N^{-1+\varepsilon}$  with arbitrarily small positive  $\varepsilon$ , which is similar to the case of classical Gaussian ensembles and the case of Wigner/sample covariance matrices. The proof of Thm. 3 will be given in Section 3.

The fact that approximate subordination holds not only for Stieltjes transforms but also for resolvents of the matrices  $H_N$  has an important consequence. Namely, this fact implies a delocalization result for eigenvectors of matrices  $H_N$ .

Let  $\{v_a^{(N)}\}_{a=1}^N$  denote an orthonormal basis of eigenvectors of  $H_N$  and let  $\lambda_a^{(N)}$  be the corresponding eigenvalues. Let  $\{e_i\}_{i=1}^N$  be the standard basis and let  $v_a^{(N)}(i)$  denote the  $i$ -th component of vector  $v_a^{(N)}$ .

**Theorem 4.** Assume that (i)  $\max \{d_L(\mu_{A_N}, \mu_\alpha), d_L(\mu_{B_N}, \mu_\beta)\} \ll N^{-1/4}$ , (ii) the pair of probability measures  $(\mu_\alpha, \mu_\beta)$  is well-behaved in an interval  $I$ , and (iii)  $\max \{\|A_N\|, \|B_N\|\} \leq K$  for all  $N$ . Then there exist constants  $C$  and  $c$  which depend only on  $(\mu_\alpha, \mu_\beta)$ ,  $I$ , and  $K$ , such that for all  $\alpha$  with  $\lambda_\alpha^{(N)} \in I$ , and all  $i = 1, \dots, N$ , it is true that

$$P \left\{ \left| v_\alpha^{(N)}(i) \right|^2 > CN^{-1/8} \log N \right\} \leq \exp \left( -cN^{1/2} \right).$$

Note that  $\sum_{i=1}^N \left| v_\alpha^{(N)}(i) \right|^2 = 1$  and the theorem says that the size of individual coordinates  $\left| v_\alpha^{(N)}(i) \right|^2$  is not greater than  $N^{-1/8}$  with high probability (up to a logarithmic factor). Hence, the result of this theorem can be interpreted as the claim that the eigenvectors of the matrix  $H_N$  cannot be localized on the set of directions with cardinality significantly smaller than  $N^{1/8}$ . It is customary to express this by saying that the delocalization length of eigenvectors is at least  $N^{1/8}$ .

The assumption that  $(\mu_\alpha, \mu_\beta)$  is well-behaved is essential in this theorem. Otherwise, we can choose  $A_N$  diagonal and  $B_N = O_N$ , the zero matrix for each  $N$ . Then  $H_N = A_N$  is diagonal and clearly the eigenvectors are localized. In this case, it is easy to calculate from system (3) that  $S_\beta(z) = 0$  and therefore the pair  $(\mu_\alpha, \mu_\beta)$  is not smooth.

The result is not optimal. The expected localization length is  $N$ , that is, eigenvalues are likely to be completely delocalized.

The question about delocalization of eigenvectors occurs in physics for eigenvectors of random Schrödinger operators on  $\mathbb{Z}^d$ . In particular, for  $d \geq 3$  it is an open problem to show the complete delocalization for small disorder. Recently, there was some progress on delocalization of eigenvectors in simpler models, namely in the case of random Wigner matrices and in the case of random band matrices. In the former case, the complete delocalization has been established recently (see [2] for a review) and the method is similar to the method that is used in this paper. In the case of band matrices, it is expected that for matrices with the band width  $W$  greater than  $\sqrt{N}$ , the case of complete delocalization holds. What was actually shown in this case is that the delocalization length is greater than  $W^{1+d/6}$  ([3]). The method is based on quantum diffusion and different from the method that is used in this paper.

The rest of the paper is organized as follows. Section 2 devoted to the proof of Theorem 2 and Proposition 1. Section 3 derives the local limit law of Theorem 3, and Section 4 is about delocalization of eigenvectors. Section 5 concludes.

## 2. APPROXIMATE SUBORDINATION

The proof of Theorem 2 and Proposition 1 will be given at the end of this section. First, we need to derive several preliminary results.

Note that if  $R_{A,N}(z)$  were equal to zero, then we could easily compute  $S_{B,N}(z)$  as

$$-(\mathbb{E}G_{H_N})^{-1} + A_N - z. \quad (6)$$

Hence, the following result is useful.

**Proposition 2.** *Let  $X$  be a random Hermitian matrix and let  $\eta > 0$ . Then all eigenvalues of the matrix  $-\left[\mathbb{E}(X - i\eta)^{-1}\right]^{-1}$  belong to the half-plane  $\mathbb{H}_\eta := \{z : \text{Im}z \geq \eta\}$ . In particular, all eigenvalues of  $-(\mathbb{E}G_{H_N})^{-1} + A_N - z$  belong to  $\mathbb{C}^+ = \{z : \text{Im}z \geq 0\}$ .*

**Proof of Proposition 2:** Every matrix  $(X(\omega) - i\eta)^{-1}$  is normal and its eigenvalues are on the border of a disc  $D_\eta$  with diameter  $[0, i/\eta]$ . (Here  $\omega$  is an element of the probability space.) Hence, by Lemma 1 below, the eigenvalues of  $\mathbb{E}(X - i\eta)^{-1}$  belong to the disc  $D_\eta$ . It follows that eigenvalues of  $-\left[\mathbb{E}(X - i\eta)^{-1}\right]^{-1}$  are in  $\mathbb{H}_\eta$ .  $\square$

**Lemma 1.** *Let  $A_j$ ,  $j = 1, \dots, m$ , be a family of normal (finite-dimensional) operators. Suppose that the eigenvalues of all  $A_j$  are contained in a disc  $D \subset \mathbb{C}$ , and let  $H = \sum p_j A_j$  be a convex combination of  $A_j$ . Then all eigenvalues of  $H$  are contained in  $D$ .*

**Proof:** By subtracting a multiple of the identity operator from all  $A_j$ , we can reduce the problem to the case when disc  $D$  has its center at 0. Assume that this is indeed the case. Let  $R$  be the radius of  $D$ . Since the operators are normal, their norms are equal to the maximum of the eigenvalue absolute values. Hence  $\|A_j\| \leq R$ . Hence,  $\|H\| \leq \sum p_j \|A_j\| \leq R$ . It follows that all eigenvalues of  $H$  have absolute value  $\leq R$ .  $\square$

It is known (see [10] or proof of Thm. 7 in [7]) that the error term in (2) can be written as follows:

$$R_{A,N} : = \frac{1}{\mathbb{E}m_{H_N}}(A_N - z)G_{A_N}(z + S_{B,N}(z))\mathbb{E}\Delta_{A_N} \quad (7)$$

$$= \frac{1}{\mathbb{E}m_{H_N}} \frac{1}{1 + (\mathbb{E}f_{B_N}/\mathbb{E}m_{H_N})G_{A_N}}\mathbb{E}\Delta_{A_N}, \quad (8)$$

where

$$\Delta_{A_N} := (m_{H_N} - \mathbb{E}m_{H_N})G_{H_N} - G_{A_N}(f_{B_N} - \mathbb{E}f_{B_N})G_{H_N}.$$

We can re-write formula (2) as:

$$\mathbb{E}G_{H_N} = G_{A_N}(z + S_{B,N}(z))(I + \Psi_{A_N}(z)),$$

where  $\Psi_{A_N} := \frac{1}{\mathbb{E}m_{H_N}}(A_N - z)\mathbb{E}\Delta_{A_N}$ .

Hence,

$$(\mathbb{E}G_{H_N})^{-1} = (I + \Psi_{A_N})^{-1}(A_N - z - S_{B,N}(z))$$

If  $Y_{A_N}$  is defined as  $(I + \Psi_{A_N})^{-1} - I$ , then we can further re-write this:

$$S_{B,N}(z) = -(\mathbb{E}G_{H_N})^{-1} + A_N - z + Y_{A_N}(A_N - z - S_{B,N}(z)), \quad (9)$$

which resembles formula (6) except for a perturbation term. We are going to estimate the size of the perturbation and to establish the following results.

**Proposition 3.** Assume that  $\max\{\|A_N\|, \|B_N\|\} \leq K$  and that  $z = E + i\eta \in \Lambda_R := \{z : z > 0, |z| \leq R\}$ . Then for all  $N \gg \eta^{-5}$ ,

$$\|Y_{A_N}(A_N - z - S_{B,N}(z))\| \leq c/(N\eta^7), \quad (10)$$

where  $c > 0$  depends only on  $K$  and  $R$ .

**Proposition 4.** Let  $A$  be an invertible diagonal matrix (not necessarily Hermitian), and  $U$  be a random Haar-distributed unitary matrix. Let  $B$  be an arbitrary matrix and assume that the matrix  $G(s) := (A + sUBU^*)^{-1}$  is bounded with probability 1 for all  $s \in [0, 1]$ . Then  $\mathbb{E}G(s)$  is diagonal for all  $s \in [0, 1]$ .

We also have the the following well-known result. (See Thm 6.3.2 on p. 365 in [5]).

**Lemma 2.** Let  $X$  be a diagonal matrix and  $A$  be an arbitrary matrix. Then for every eigenvalue  $\hat{\lambda}_i$  of  $X + A$ , there exists an eigenvalue  $\lambda_i$  of  $X$  such that  $|\hat{\lambda}_i - \lambda_i| \leq \|A\|$ .

Assuming the validity of these results it is easy to prove Theorem 2:

**Proof of Theorem 2:** Take the basis in which  $A_N$  is diagonal. From Propositions 2 and 4 we conclude that  $-(\mathbb{E}G_{H_N})^{-1} + A_N - z$  is diagonal and its eigenvalues belong to the halfplane  $\mathbb{C}^+ = \{z : \text{Im}z \geq 0\}$ . From Proposition 3,

$$\|Y_{A_N}(A_N - z - S_{B,N}(z))\| \leq c/(N\eta^7).$$

Hence, formula (9) and Lemma 2 imply that

$$\text{Im}(S_{B,N}(z)) \geq -\frac{c}{N\eta^7}.$$

The proof for  $\text{Im}(S_{A,N}(z))$  is similar.  $\square$

For the proof of Prop. 3 we need several estimates which we formulate as Lemmas. In all of them we assume that  $\max\{\|A_N\|, \|B_N\|\} \leq K$  and  $z = E + i\eta \in \Lambda_R := \{z : z > 0, |z| \leq R\}$ .

**Lemma 3.** Let  $X$  be a random Hermitian matrix,  $\|X\| \leq K$ . Then  $\text{Im}\mathbb{E}\left[\frac{1}{N}\text{Tr}(X - z)^{-1}\right] \geq c\eta$ , where  $c$  depends on  $K$  and  $R$  only. In particular,  $\left|(\mathbb{E}m_{H_N}(z))^{-1}\right| \leq c'/\eta$ , where  $c'$  depends only on  $K$  and  $R$ .

**Proof:** We can write

$$\text{Im}\mathbb{E}\left[\frac{1}{N}\text{Tr}\frac{1}{X - a - i\eta}\right] = \mathbb{E}\left[\frac{1}{N}\text{Tr}\frac{\eta}{(X - a)^2 + \eta^2}\right].$$

Since  $(X - a)^2 + \eta^2 I \leq ((K + R)^2 + R^2) I$ , hence  $((X - a)^2 + \eta^2 I)^{-1} \geq ((K + R)^2 + R^2)^{-1} I$  and we conclude that

$$\mathbb{E}\left[\frac{1}{N}\text{Tr}\frac{\eta}{(X - a)^2 + \eta^2}\right] \geq c\eta.$$

$\square$

**Lemma 4.** For a numeric  $c > 0$  and every  $\delta > 0$ ,

$$P \{ |m_{H_N}(z) - \mathbb{E}m_{H_N}(z)| > \delta \} \leq \exp \left( -\frac{c\delta^2\eta^4}{\|B_N\|^2} N^2 \right), \quad (11)$$

and

$$P \{ |f_{B_N}(z) - \mathbb{E}f_{B_N}(z)| > \delta \} \leq \exp \left[ -\frac{c\delta^2\eta^4}{\|B_N\|^4} N^2 / \left( 1 + \frac{\eta}{\|B_N\|} \right)^2 \right]. \quad (12)$$

For the proof see Prop. 4 in [7].

**Lemma 5.**

$$P \{ \|\Delta_A(z)\| \geq \varepsilon \} \leq \exp [-c\varepsilon^2\eta^8 N^2],$$

and  $\|\mathbb{E}\Delta_A(z)\| \leq c/(N\eta^4)$  where constants are positive and depend only on  $K$  and  $R$ .

**Proof:** For any Hermitian matrix  $X$  and  $\eta > 0$  it is true that  $\|(X - i\eta)^{-1}\| \leq 1/\eta$ . Hence, by using Lemma 4 we can write

$$\begin{aligned} P \{ \|(m_{H_N}(z) - \mathbb{E}m_{H_N}(z)) G_{H_N}\| \geq \delta/\eta \} &\leq \exp \left[ -c \frac{\delta^2\eta^4}{\|B_N\|^2} N^2 \right], \\ P \{ \|G_{A_N}(f_{B_N}(z) - \mathbb{E}f_{B_N}(z)) G_{H_N}\| \geq \delta/\eta^2 \} &\leq \exp \left[ -c \frac{\delta^2\eta^4}{\|B_N\|^4} N^2 / \left( 1 + \frac{\eta}{\|B_N\|} \right)^2 \right]. \end{aligned}$$

Set  $\varepsilon = \delta/\eta$  and  $\varepsilon = \delta/\eta^2$  in the first and the second inequalities, respectively, and use the triangle inequality for norms and we will get that

$$\begin{aligned} P \{ \|\Delta_A(z)\| \geq \varepsilon \} &\leq \exp \left[ -\frac{c\varepsilon^2 N^2}{\|B_N\|^2} \min \left\{ \eta^6, \eta^8 / \left( 1 + \frac{\eta}{\|B_N\|} \right)^2 \right\} \right] \\ &\leq \exp [-c\varepsilon^2\eta^8 N^2], \end{aligned}$$

where we used the assumption that  $|z| \leq R$  and where the constant  $c$  in the last inequality can depend on  $\|B_N\|$  and  $R$ . For the second claim, note that  $\|\mathbb{E}\Delta_A\| \leq \mathbb{E}\|\Delta_A\|$  by the convexity of norm, and  $\mathbb{E}\|\Delta_A\|$  can be estimated by using the first claim of Lemma 5 and the equality

$$\mathbb{E}X = \int_0^\infty (1 - \mathcal{F}_X(t)) dt,$$

valid for every positive random variable  $X$  and its cumulative distribution function  $\mathcal{F}_X(t)$ .

In our case, we obtain

$$\mathbb{E}\|\Delta_A\| \leq \int_0^\infty \exp [-ct^2\eta^8 N^2] dt = \frac{c'}{N\eta^4}.$$

□

**Lemma 6.**  $\|\Psi_{A_N}\| \leq c/(N\eta^5)$ , where  $c > 0$  depends only on  $K$  and  $R$ .

**Proof:** This is a direct consequence of Lemmas 3 and 5, which estimate  $\left| (\mathbb{E}m_{H_N})^{-1} \right|$  and  $\|\mathbb{E}\Delta_{A_N}\|$  as  $c/\eta$  and  $c/(N\eta^4)$ , respectively. □

**Lemma 7.** *If  $\|X\| \leq \varepsilon < 1/2$ , then  $\|(I + X)^{-1} - I\| \leq 2\varepsilon$ . In particular, for all  $N \gg \eta^{-5}$ ,  $\|Y_{A_N}\| \leq c/(N\eta^5)$ , where  $c > 0$  depends only on  $K$  and  $R$ .*

**Proof:**  $(I + X)^{-1} - I = X \sum_{k=0}^{\infty} (-1)^k X^k$ , and  $\left\| -X \sum_{k=0}^{\infty} (-1)^k X^k \right\| \leq \varepsilon \sum_{k=0}^{\infty} \varepsilon^k < 2\varepsilon$ . The second claim follows from Lemma 6.  $\square$

**Lemma 8.**  $|S_{B,N}(z)| \leq c/\eta^2$ .

**Proof:** By definition,  $S_{B,N}(z) = -\mathbb{E}f_{B_N}(z) / \mathbb{E}m_{H_N}(z)$ . From Lemma 3,  $\left| (\mathbb{E}m_{H_N}(z))^{-1} \right| < c/\eta$ . In addition,

$$|\mathbb{E}f_{B_N}(z)| = \left| \mathbb{E} \frac{1}{N} \text{Tr} \left( B_N \frac{1}{H_N - z} \right) \right| \leq \|B_N\| \mathbb{E} \left( \left\| \frac{1}{H_N - z} \right\| \right) \leq c \frac{1}{\eta}.$$

$\square$

**Proof of Prop. 3:** Estimates in Lemmas 7 and 8 imply (10).  $\square$

**Proof of Prop. 4:** For proof we need several Lemmas.

Let  $u_{ij}$  and  $u_{ij}^*$  denote the matrix elements of random matrices  $U$  and  $U^*$ , respectively, where  $U$  is distributed uniformly on the unitary group  $\mathcal{U}(N)$ . Consider the product

$$\Pi = u_{i_1 j_1} u_{k_1 l_1}^* u_{i_2 j_2} u_{k_2 l_2}^* \dots u_{i_n j_n} u_{k_n l_n}^*.$$

Let  $L(i)$  be the number of times index  $i$  appears as the first index of  $u_{i_s j_s}$ , and let  $R(i)$  be the number of times it appears as the second index in  $u_{k_s l_s}^*$ .

**Lemma 9.** *If  $L(i) \neq R(i)$ , then  $\mathbb{E}\Pi = 0$ .*

**Proof:** Let  $\Delta_t(i)$  be a diagonal matrix that coincides with the identity matrix everywhere except at the  $i$ -th place of the main diagonal where it has  $\exp(t\sqrt{-1})$  instead of 1. Assume that  $t$  is random and distributed uniformly on  $[0, 2\pi]$ . For every  $t$ ,  $\Delta_t U$  is distributed uniformly on  $\mathcal{U}(N)$ , hence if  $\Pi_t$  denotes the product analogous to  $\Pi$  but computed for the matrix  $\Delta_t U$ , then the conditional expectation  $\mathbb{E}(\Pi_t | t) = \mathbb{E}(\Pi)$ . Hence  $\mathbb{E}(\Pi_t) = \mathbb{E}(\Pi)$ , where the first expectation is taken both over  $U$  and  $t$ . However, if  $L(i) \neq R(i)$ , then we have  $\Pi_t = e^{kt\sqrt{-1}} \Pi$  with  $k \neq 0$ . Hence  $\int_0^{2\pi} \Pi_t dt = 0$  and  $\mathbb{E}(\Pi_t) = 0$ .  $\square$

**Lemma 10.** *Let  $A$  is a diagonal matrix (not necessarily Hermitian), and  $B$  is arbitrary. Let  $U$  be a random Haar-distributed unitary matrix. Let  $n$  be a non-negative integer. Then*

$$\mathbb{E}[(UBU^*A)^n UBU^*]$$

*is diagonal.*

**Proof:** A typical matrix element of  $(UBU^*A)^n UBU^*$  is

$$u_{i_1 j_1} b_{j_1 k_1} u_{k_1 l_1}^* a_{l_1 i_2} u_{i_2 j_2} b_{j_2 k_2} u_{k_2 l_2}^* \dots u_{i_{n+1} j_{n+1}} b_{j_{n+1} k_{n+1}} u_{k_{n+1} l_{n+1}}^*.$$

We need to show that the expectation of this expression is zero if  $i_1 \neq l_{n+1}$ .

By assumption,  $A$  is diagonal, hence every non-zero product has to satisfy the equality  $l_s = i_{s+1}$ . Let  $i_1 = i$ . Then  $L(i) \geq R(i)$  since if  $i$  appears as the second index of  $u_{k_s l_s}^*$ ,  $s \leq n$ , then it must appear as the first index of  $u_{i_{s+1} j_{s+1}}$ , and it also appears as the first index of  $u_{i_1 j_1}$ . Moreover, the equality holds if and only if  $i$  appears as the second index of  $u_{k_{n+1} l_{n+1}}^*$ . By Lemma 9, the expectation of the product is zero unless  $L(i) = R(i)$ , that is, unless  $i_1 = l_{n+1}$ .  $\square$

Now we can complete the proof of Prop. 4. Let

$$G(s) = (A + sUBU^*)^{-1}.$$

If  $s$  is small, then we can expand  $G(s)$  as a convergent power series in  $s$ ,

$$G(s) = A^{-1} + A^{-1} \left( \sum_{n=0}^{\infty} (UBU^* A^{-1})^n UBU^* s^{n+1} \right) A^{-1}.$$

By using Lemma 10, we find that if  $s$  is sufficiently small and the series are convergent, then  $\mathbb{E}G(s)$  is diagonal, that is, that all off-diagonal elements of  $\mathbb{E}G(s)$  are zero. By assumption the matrix  $G(s) := (A + sUBU^*)^{-1}$  is bounded with probability 1 for all  $s \in [0, 1]$ . Hence,  $G(s)$  can be analytically continued along this interval, and  $\mathbb{E}G(s)$  is an analytic function for all  $s \in [0, 1]$ . Off-diagonal matrix elements of  $\mathbb{E}G(s)$  are zero for all sufficiently small  $s$ , therefore they are zero for all  $s \in [0, 1]$  by properties of complex-analytic functions. Hence,  $\mathbb{E}G(s)$  is diagonal for all  $s \in [0, 1]$ .  $\square$

**Proof of Proposition 1:** If  $A_N = B_N$ , then Theorem 1 implies that

$$\operatorname{Im} z' = \operatorname{Im}(z + S_{B,N}(z)) \geq \eta,$$

where  $\eta := \operatorname{Im} z$ . Similarly, if  $A \neq B$  and  $N \geq c/\eta^8$  for sufficiently large  $c$ , then Theorem 2 implies that

$$\operatorname{Im} z' = \operatorname{Im}(z + S_{B,N}(z)) \geq \eta/2.$$

Hence, in both cases  $\|G_{A_N}(z')\| \leq c/\eta$ . Since  $R_{A,N}(z) = G_{A_N}(z') \Psi_{A_N}(z)$ , we can use Lemma 6 in order to obtain

$$\|R_{A,N}(z)\| \leq \|G_{A_N}(z')\| \|\Psi_{A_N}(z)\| \leq \frac{c}{N\eta^6}.$$

$\square$

### 3. LOCAL LAW FOR EIGENVALUES

We will prove Theorem 3 at the end of this section after several preliminary results.

First, we can exclude  $m_{\boxplus}(z)$  from system (3):

$$\begin{aligned} m_{\alpha}(z + S_{\beta}(z)) + \frac{1}{z + S_{\alpha}(z) + S_{\beta}(z)} &= 0, \\ m_{\beta}(z + S_{\alpha}(z)) + \frac{1}{z + S_{\alpha}(z) + S_{\beta}(z)} &= 0. \end{aligned} \tag{13}$$

A similar system can be written for  $S_{A,N}(z)$  and  $S_{B,N}(z)$ :

$$\begin{aligned} m_{A_N}(z + S_{B,N}(z)) + \frac{1}{z + S_{A,N}(z) + S_{B,N}(z)} &= r_{A,N}(z), \\ m_{B_N}(z + S_{A,N}(z)) + \frac{1}{z + S_{A,N}(z) + S_{B,N}(z)} &= r_{B,N}(z), \end{aligned} \quad (14)$$

where  $r_{A,N}(z) = N^{-1} \text{Tr}(R_{A,N}(z))$ ,  $r_{B,N}(z) = N^{-1} \text{Tr}(R_{B,N}(z))$ , and  $R_{A,N}, R_{B,N}$  are defined as in (7).

There is also an “intermediate” system which uses  $m_\alpha$  and  $m_\beta$  as in (13) and has non-zero right hand side as in (14):

$$\begin{aligned} m_\alpha(z + \tilde{S}_\beta(z)) + \frac{1}{z + \tilde{S}_\alpha(z) + \tilde{S}_\beta(z)} &= r_{A,N}(z), \\ m_\beta(z + \tilde{S}_\alpha(z)) + \frac{1}{z + \tilde{S}_\alpha(z) + \tilde{S}_\beta(z)} &= r_{B,N}(z). \end{aligned} \quad (15)$$

We are going to show that the solutions of (15)  $(\tilde{S}_\alpha(z), \tilde{S}_\beta(z))$  are close to solutions of both (13)  $(S_\alpha(z), S_\beta(z))$  and (14)  $(S_{A,N}(z), S_{B,N}(z))$ . In this way, we will show that the solutions of (13) and (14) are close to each other.

**Proposition 5.** *Assume that (i) a pair of probability measures  $(\mu_\alpha, \mu_\beta)$  is well-behaved at  $E$ , (ii)  $\max\{\|A_N\|, \|B_N\|\} \leq K$ , (iii)  $z = E + i\eta \in \{z : \eta > 0, |z| \leq R\}$ , and (iv)  $N\eta^8 > c_1$ , where  $c_1$  depends on  $K$  and  $R$ . Then for some positive  $s_0, \eta_0$  and  $c$ , if  $d_L(\mu_A, \mu_\alpha) < s \leq s_0$  and  $d_L(\mu_B, \mu_\beta) < s \leq s_0$ , then*

$$|S_\alpha(E + i\eta) - S_{A,N}(E + i\eta)| < \frac{c}{N\eta^6} + cs,$$

and

$$|S_\beta(E + i\eta) - S_{B,N}(E + i\eta)| < \frac{c}{N\eta^6} + cs.$$

for all  $\eta < \eta_0$ .

**Proof:** The claim is the direct consequence of Lemmas 11 and 12 below, in which the differences between solutions of (13) and (15), and between solutions of (15) and (14) are estimated.  $\square$

In all lemmas below we assume that a pair of probability measures  $(\mu_\alpha, \mu_\beta)$  is well-behaved at  $E$ , that  $\max\{\|A_N\|, \|B_N\|\} \leq K$ , and that  $z = E + i\eta \in \{z : \eta > 0, |z| \leq R\}$ ,

**Lemma 11.** *Assume that  $N\eta^8 > c_1$ , where  $c_1$  is a constant that may depend on  $K$  and  $R$  and the pair of measures  $(\mu_\alpha, \mu_\beta)$ . Then for some positive  $\eta_0$  and  $c$ , it is true that*

$$\left| \tilde{S}_\alpha(E + i\eta) - S_\alpha(E + i\eta) \right| < \frac{c}{N\eta^6},$$

and

$$\left| \tilde{S}_\beta(E + i\eta) - S_\beta(E + i\eta) \right| < \frac{c}{N\eta^6},$$

for all  $\eta < \eta_0$ .

**Proof:** Let  $F(t)$  be a function  $\mathbb{C}^2 \rightarrow \mathbb{C}^2$  defined by the formula:

$$F : \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} \rightarrow \begin{pmatrix} (z + s_1 + s_2)^{-1} + m_\alpha(z + s_2) - r_{A,N}(z) \\ (z + s_1 + s_2)^{-1} + m_\beta(z + s_1) - r_{B,N}(z) \end{pmatrix}.$$

Let us use the norm  $\|(x_1, x_2)\| = (|x_1|^2 + |x_2|^2)^{1/2}$ .

The system (15) can be written as  $F(x) = 0$ . By Proposition 1.,  $\|F(S_\alpha, S_\beta)\| = \|(r_{A,N}, r_{B,N})\| \leq c/(N\eta^6)$  provided that  $N\eta^8 \geq c_1$ .

The derivative of  $F$  with respect to  $s$  is

$$F' = \begin{pmatrix} -(z + s_1 + s_2)^{-2} & -(z + s_1 + s_2)^{-2} + m'_\alpha(z + s_2) \\ -(z + s_1 + s_2)^{-2} + m'_\beta(z + s_1) & -(z + s_1 + s_2)^{-2} \end{pmatrix}.$$

The determinant of this matrix is

$$[m'_\alpha(z + s_2) + m'_\beta(z + s_1)](z + s_1 + s_2)^{-2} - m'_\alpha(z + s_2)m'_\beta(z + s_1).$$

This expression is non-zero at  $(S_\alpha, S_\beta)$  by assumption (5). This, together with another application of the assumption of smoothness, shows that the entries of the matrix  $[F']^{-1}$  are bounded at  $(S_\alpha(z), S_\beta(z))$  for all  $z$  in a neighborhood of  $E$ . This shows that operator norm of  $[F']^{-1}$  is bounded.

Similarly, the assumption of smoothness implies that the operator norm of  $F''$  is bounded for all  $(s_1, s_2)$  in a neighborhood of  $(S_\alpha(z), S_\beta(z))$  and all values of parameter  $z$  in a neighborhood of  $E$ .

It follows by the Newton-Kantorovich theorem ([6]) that the solution of the equation  $F(s) = 0$ , which is  $(\tilde{S}_\alpha(z), \tilde{S}_\beta(z))$  exists provided that  $N\eta^8 \geq c_1$  and satisfies the inequalities:

$$|\tilde{S}_\alpha(z) - S_\alpha(z)| < c/(N\eta^6),$$

and

$$|\tilde{S}_\beta(z) - S_\beta(z)| < c/(N\eta^6).$$

□

**Corollary 1.** *If  $N \gg \eta^{-8}$  and  $\eta \ll 1$ , then  $\tilde{S}_\alpha(E) := \lim_{\eta \downarrow 0} \tilde{S}_\alpha(E + i\eta)$  and  $\tilde{S}_\beta(E) := \lim_{\eta \downarrow 0} \tilde{S}_\beta(E + i\eta)$  exist and both  $\text{Im}\tilde{S}_\alpha(E)$  and  $\text{Im}\tilde{S}_\beta(E)$  are strictly positive. Moreover,*

$$\frac{1}{m'_\alpha(E + \tilde{S}_\beta(E))} + \frac{1}{m'_\beta(E + \tilde{S}_\alpha(E))} - (E + \tilde{S}_\alpha(E) + \tilde{S}_\beta(E))^2 \neq 0.$$

**Lemma 12.** *Assume that  $N\eta^8 > c_1$ , where  $c_1$  is a constant that may depend on  $K$  and  $R$  and the pair of measures  $(\mu_\alpha, \mu_\beta)$ . Then for some positive  $s_0$ ,  $\eta_0$  and  $c$ , if  $d_L(\mu_A, \mu_\alpha) < s \leq s_0$  and  $d_L(\mu_B, \mu_\beta) < s \leq s_0$ , then*

$$|\tilde{S}_\alpha(E + i\eta) - S_{A,N}(E + i\eta)| < cs,$$

and

$$|\tilde{S}_\beta(E + i\eta) - S_{B,N}(E + i\eta)| < cs,$$

for all  $\eta < \eta_0$ .

**Proof:** We proceed similarly to the proof of Lemma 11 and write system (14) as  $F_N(x) = 0$ , where

$$F : \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} \rightarrow \begin{pmatrix} (z + s_1 + s_2)^{-1} + m_{A_N}(z + s_2) - r_{A,N}(z) \\ (z + s_1 + s_2)^{-1} + m_{B,N}(z + s_1) - r_{B,N}(z) \end{pmatrix}.$$

Corollary 1 implies that  $[F']^{-1}$  is not 0 at  $(\tilde{S}_\alpha, \tilde{S}_\beta)$ . It is also possible to check that  $F''$  is bounded in a neighborhood of  $(\tilde{S}_\alpha, \tilde{S}_\beta)$ . Then we use the Newton-Kantorovich theorem to find a solution of (14) by starting with the initial guess  $(\tilde{S}_\alpha, \tilde{S}_\beta)$ . In order to estimate the initial error, we can use the following lemmas.

**Lemma 13.** *let  $m_1(z)$  and  $m_2(z)$  denote the Stieltjes transforms of measures  $\mu_1$  and  $\mu_2$ , respectively. Let  $d_L(\mu_1, \mu_2) = s$  and  $z = E + i\eta$ , where  $\eta > 0$ . Then,*

- (a)  $|m_1(z) - m_2(z)| < cs\eta^{-1} \max\{1, \eta^{-1}\}$  where  $c > 0$  is a numeric constant, and
- (b)  $|\frac{d^r}{dz^r}(m_1(z) - m_2(z))| < c_r s \eta^{-1-r} \max\{1, \eta^{-1}\}$  where  $c > 0$  are numeric constants.

This lemma was proved as Lemma 2.2 in [8].

**Lemma 14.** *Suppose that  $d_L(\mu_\alpha, \mu_{A_N}) < s$  and  $d_L(\mu_\beta, \mu_{B_N}) < s$ . Then,*

$$\left| \frac{1}{z + \tilde{S}_\alpha(z) + \tilde{S}_\beta(z)} + m_{A_N}(z + \tilde{S}_\beta(z)) - r_{A,N}(z) \right| \leq cs,$$

and

$$\left| \frac{1}{z + \tilde{S}_\alpha(z) + \tilde{S}_\beta(z)} + m_{B_N}(z + \tilde{S}_\alpha(z)) - r_{B,N}(z) \right| \leq cs,$$

for all  $\eta > cN^{-1/8}$ , where  $c$  is a positive constant that depends only on the pair of measures  $\mu_\alpha$  and  $\mu_\beta$ .

**Proof:**  $\tilde{S}_\alpha(z)$  and  $\tilde{S}_\beta(z)$  satisfy the equations of system (15), which implies that it is enough to show that

$$\left| m_{A_N}(z + \tilde{S}_\beta(z)) - m_\alpha(z + \tilde{S}_\beta(z)) \right| < cs,$$

and

$$\left| m_{B_N}(z + \tilde{S}_\alpha(z)) - m_\beta(z + \tilde{S}_\alpha(z)) \right| < cs,$$

This follows from Corollary 1 and Lemma 13. Indeed, let  $\eta_0(E) > 0$  denote  $\min\{\text{Im}\tilde{S}_\alpha(E), \text{Im}\tilde{S}_\beta(E)\}$ . Then, by Lemma 13,

$$\left| m_{A_N}(z + \tilde{S}_\beta(z)) - m_\alpha(z + \tilde{S}_\beta(z)) \right| < cs \min\left\{ \frac{1}{\eta_0(E)}, \frac{1}{\eta_0^2(E)} \right\},$$

and similar estimate holds for the difference  $\left| m_{B_N}(z + \tilde{S}_\alpha(z)) - m_\beta(z + \tilde{S}_\alpha(z)) \right|$ .  $\square$

By using this Lemma and the Newton-Kantorovich theorem we find that the distance between  $(\tilde{S}_\alpha, \tilde{S}_\beta)$  and the solution of (14) is smaller than  $cs$ . This completes the proof of Proposition 5.  $\square$ .

**Proposition 6.** *Let the assumptions of Prop. 5 hold. Then for some positive  $s_0, \eta_0$  and  $c$ , if  $d_L(\mu_A, \mu_\alpha) < s \leq s_0$  and  $d_L(\mu_B, \mu_\beta) < s \leq s_0$ , then*

$$|\mathbb{E}m_{H_N}(E + i\eta) - m_\boxplus(E + i\eta)| < \frac{c}{N\eta^6} + cs,$$

for all  $\eta < \eta_0$ .

**Proof:** Since  $\mathbb{E}m_{H_N} = (z + S_{A,N} + S_{B,N})^{-1}$  and  $m_\boxplus = (z + S_\alpha + S_\beta)^{-1}$ , therefore

$$\mathbb{E}m_{H_N} - m_\boxplus = \frac{1}{z + S_{A,N} + S_{B,N}} - \frac{1}{z + S_\alpha + S_\beta} (S_\alpha + S_\beta - S_{A,N} - S_{B,N}).$$

The denominator is bounded away from zero if  $N\eta^8 > c_1$  by Prop. 5 ( $\text{Im}S_\alpha(E)$  and  $\text{Im}S_\beta(E)$  are bounded away from 0 by the assumption of Prop. 5 and  $S_{A,N}$  and  $S_{B,N}$  are close to  $\text{Im}S_\alpha(E)$  and  $\text{Im}S_\beta(E)$ , respectively, by its conclusion). The numerator can also be estimated by Proposition 5.  $\square$

**Proposition 7.** *Suppose that the assumptions of Proposition 5 hold. Let  $\eta^* = M\eta$  and  $I_{\eta^*} = [E - \eta^* + i\eta, E + \eta^* + i\eta]$ . Then for some positive  $c$  and  $c_1$ , that can depend on  $K, R, M$ ,*

$$P \left\{ \sup_{z \in I_{\eta^*}} |m_{H,N}(z) - m_\boxplus(z)| > \frac{c}{N^{1/4}} + cs \right\} \leq \exp(-c_1 (\log N)^2), \quad (16)$$

if  $\eta < \eta_0$ .

**Proof:** This is a combination of statements in Proposition 6 and in the following strengthening of Lemma 4, which was proved as Corollary 6 in [7].

**Lemma 15.** *For some positive  $c$  and  $c_1$  which can depend on  $M$  and for all  $\delta > 0$ ,*

$$P \left\{ \sup_{z \in I_{\eta^*}} |m_{H_N}(z) - \mathbb{E}m_{H_N}(z)| > \delta \right\} \leq \exp \left( -\frac{c\delta^2\eta^4}{\|B_N\|^2 N^2} \right),$$

provided that  $N \geq c_1 \left( \sqrt{-\log(\eta\delta)} \right) / (\eta^2\delta)$ .

Let us take  $\delta = c \log N / (N\eta^2)$ . Then the assumption of Lemma 15 is satisfied provided that  $\eta = O(N^{-1+\varepsilon})$ . In particular, if  $\eta = cN^{-1/8}$  for a sufficiently large  $c$ , then Lemma 15 implies that

$$P \left\{ \sup_{z \in I_{\eta^*}} |m_H(z) - \mathbb{E}m_H(z)| > \frac{\log N}{N^{3/4}} \right\} \leq \exp(-c (\log N)^2).$$

In addition, Proposition 6 says that

$$|\mathbb{E}m_H(z) - m_\boxplus(z)| < \frac{c}{N^{1/4}} + cs.$$

Together, these statements imply the claim of Proposition 7.  $\square$

**Proof of Theorem 3:** The proof is similar to the proof of Corollary 4.2 (local semicircle law) in [4]. We provide details for the convenience of the reader. Define  $\eta = N^{-1/8}$ , and  $\eta^* = M\eta$ . Let

$$\begin{aligned} R(\lambda) &: = \frac{1}{\pi} \int_{E-\eta^*}^{E+\eta^*} \frac{\eta}{(x-\lambda)^2 + \eta^2} dx \\ &= \frac{1}{\pi} \left( \arctan \left( \frac{E-\lambda}{\eta} + M \right) - \arctan \left( \frac{E-\lambda}{\eta} - M \right) \right). \end{aligned}$$

Then  $R = 1_{I^*} + T_1 + T_2 + T_3$ , where  $T_1, T_2, T_3$  satisfy the following properties:

$$|T_1| \leq c/\sqrt{M}, \quad \text{supp}(T_1) \subset I_1 = [E - 2\eta^*, E + 2\eta^*],$$

$$|T_2| \leq 1, \quad \text{supp}(T_2) \subset J_1 \cup J_2,$$

where  $J_1$  and  $J_2$  are intervals of length  $\sqrt{M}\eta$  with midpoints at  $E - \eta^*$  and  $E + \eta^*$ , respectively,

$$|T_3| \leq \frac{C\eta\eta^*}{(\lambda - E)^2 + (\eta^*)^2}, \quad \text{supp}(T_3) \in I_1^c.$$

Note that

$$\begin{aligned} \frac{N_{\eta^*}(E)}{2\eta^*N} &= \frac{1}{2\eta^*} \int 1_{I^*}(\lambda) \mu_{H_N}(d\lambda) \\ &= \frac{1}{2\eta^*} \int R(\lambda) \mu_{H_N}(d\lambda) - \frac{1}{2\eta^*} \int (T_1 + T_2 + T_3) \mu_{H_N}(d\lambda). \end{aligned}$$

The last integral can be estimated as  $O(M^{-1/2})$ . For the main term we have

$$\frac{1}{2\eta^*} \int_{I^*} \frac{1}{\pi} \text{Im} m_{\boxplus}(x + i\eta) dx + \frac{1}{2\eta^*} \int_{I^*} \frac{1}{\pi} \text{Im} (m_{H_N}(x + i\eta) - m_{\boxplus}(x + i\eta)) dx.$$

The first part converges to  $\rho_{\boxplus}(E)$  because the assumption that  $(\mu_\alpha, \mu_\beta)$  is well-behaved at  $E$  implies that  $\mu_\alpha \boxplus \mu_\beta$  has an analytic density in a neighborhood of  $E$ . For the second term we can use the estimate in Proposition 7 that shows that this term converges to 0 in probability as  $N \rightarrow \infty$ .  $\square$

#### 4. DELOCALIZATION

**Proof of Theorem 4:** For simplicity of notation, let us omit the subscript  $H_N$  in  $G_{H_N}$ , i.e.,  $G := G_{H_N}$ . We can write

$$\text{Im} G_{ii}(E + i\eta) = \sum_{a=1}^N \frac{\eta |v_a(i)|^2}{(\lambda_a - E)^2 + \eta^2}.$$

Let us set  $E = \lambda_a$ , then

$$|G_{ii}(E + i\eta)| \geq \text{Im} G_{ii}(E + i\eta) \geq \frac{|v_a(i)|^2}{\eta},$$

and therefore

$$|v_a(i)|^2 \leq \eta |G_{ii}(E + i\eta)|.$$

By using Lemma 18 below with  $\eta = N^{-1/8} \log N$  we obtain the claim of the theorem.  $\square$

**Lemma 16.** *Let  $1 \gg \eta \gg N^{-1/8}$ , and assume that (i)  $\max \{d_L(\mu_{A_N}, \mu_\alpha), d_L(\mu_{B_N}, \mu_\beta)\} \ll N^{-1/4}$ , (ii) the pair of probability measures  $(\mu_\alpha, \mu_\beta)$  is well-behaved at  $E$ , and (iii)  $\max \{\|A_N\|, \|B_N\|\} \leq K$  for all  $N$ . Then, for all  $k = 1, \dots, N$ ,*

$$|\mathbb{E}G_{kk}(E + i\eta)| \leq C,$$

where  $C$  is a constant that depends only on  $K$ ,  $E$ , and the pair  $(\mu_\alpha, \mu_\beta)$ .

**Proof:** Recall that (as in (2))

$$\mathbb{E}G_{H_N}(z) = G_{A_N}(z + S_{B,N}(z)) + R_{A,N}(z), \quad (17)$$

By Proposition 5, assumptions of the lemma imply that

$$|S_\beta(E + i\eta) - S_{B,N}(E + i\eta)| < \frac{c}{N\eta^6} + cs,$$

where  $s = \max \{d_L(\mu_{A_N}, \mu_\alpha), d_L(\mu_{B_N}, \mu_\beta)\}$ . Hence  $z - S_{B,N}(z)$  has a positive imaginary part for sufficiently large  $N$ . In addition, by Proposition 1,

$$\|R_{A,N}(z)\| \leq c/(N\eta^6).$$

Hence,

$$\|\mathbb{E}G_{H_N}(E + i\eta)\| \leq C,$$

which implies the claim of the lemma.  $\square$

**Lemma 17.** *Let  $z = E + i\eta$  where  $\eta > 0$ . Then, for a numeric  $c > 0$  and every  $\delta > 0$ ,*

$$P\{|G_{ii}(z) - \mathbb{E}G_{ii}(z)| > \delta\} \leq \exp\left(-\frac{c\delta^2\eta^4}{\|B_N\|^2}N\right), \quad (18)$$

**Proof:** By proceeding in the same fashion as in the proof of Lemma 4, we compute

$$\frac{\partial G_{ii}}{\partial B_{xy}} = -G_{ix}G_{yi}.$$

Since  $G$  is symmetric, this matrix has rank 1 and its only eigenvalue equals  $\sum_a (G_{ai})^2$ . Therefore,

$$\left\|\frac{\partial G_{ii}}{\partial B_{xy}}\right\|_2 = \left|\sum_a (G_{ai})^2\right| \leq \sum_a |G_{ai}|^2 = \|Ge_i\|^2 \leq \|G\|^2 \leq \frac{1}{\eta^2}.$$

Hence, we can estimate

$$\|d_X G_{ii}\|_2 \leq \frac{2\|B_N\|}{\eta^2}.$$

The rest of the proof is similar to the proof of Lemma 4.  $\square$

**Lemma 18.** *Let  $z = E + i\eta$ ,  $1 \gg \eta \gg N^{-1/8}$ , and assume that (i)  $\max \{d_L(\mu_{A_N}, \mu_\alpha), d_L(\mu_{B_N}, \mu_\beta)\} \ll N^{-1/4}$ , (ii) the pair of probability measures  $(\mu_\alpha, \mu_\beta)$  is well-behaved at  $E$ , and (iii)  $\max \{\|A_N\|, \|B_N\|\} \leq K$  for all  $N$ . Then there exists  $C > 0$  such that for all  $\delta > 0$ :*

$$P\{|G_{ii}(z)| \geq C + \delta\} \leq \exp\left(-c\delta^2 N^{1/2}\right).$$

**Proof:** This is a direct consequence of Lemmas 16 and 17. (In Lemma 17 we use  $\eta = cN^{-1/8}$ .)  $\square$

## 5. CONCLUDING REMARKS

We showed that there exists a function  $S_{B,N}(z)$  that have the following properties:

- (i)  $S_{B,N}(z)$  is analytical in  $\mathbb{C}^+$ ;
- (ii) If  $\text{Im}z \geq cN^{-1/5}$ , and  $N$  is sufficiently large then

$$\text{Im}S_{B,N}(z) \geq -\frac{c}{N(\text{Im}z)^7}.$$

Moreover, if  $\mu_{A_N} = \mu_{B_N}$ , then  $\text{Im}S_{B,N}(z) \geq 0$  for all  $z$  with  $\text{Im}z > 0$ , and therefore  $S_{B,N}$  maps  $\mathbb{C}^+$  to  $\mathbb{C}^+$ ;

- (iii) if  $R_{A,N}(z)$  is defined as

$$R_{A,N}(z) = G_{H_N}(z) - G_{A_N}(z + S_{B,N}(z)),$$

then  $\|R_{A,N}(z)\| \rightarrow 0$  provided that  $\text{Im}z \geq cN^{-1/8}$ .

One remaining question is whether it is true that  $S_{B,N}(z)$  maps  $\mathbb{C}^+$  to  $\mathbb{C}^+$  without the additional assumption that  $\mu_{A_N} = \mu_{B_N}$ .

By using this result, we showed that a local limit law holds for the empirical eigenvalue measure  $\mu_{H_N}$  with the window length  $N^{-1/8}$ . This result improves over the result in [7]. However, it is still far from the expected optimal result with the window length  $N^{-1+\varepsilon}$ .

Finally, we used the subordination property to show that the localization length of eigenvectors is greater than  $N^{1/8}$ . Again, this result is not optimal and the expected localization length is  $O(N)$  (complete delocalization).

## REFERENCES

- [1] Philippe Biane. Processes with free increments. *Mathematische Zeitschrift*, 227:143–174, 1998.
- [2] Laszlo Erdos. Universality of wigner random matrices: a survey of recent results. [arxiv.org/abs/1004.0861](http://arxiv.org/abs/1004.0861), 2010.
- [3] Laszlo Erdos and Antti Knowles. Quantum diffusion and eigenfunction delocalization in a random band matrix model. *Communications in Mathematical Physics*, 303:509–554, 2011.
- [4] Laszlo Erdos, Benjamin Schlein, and Horng-Tzer Yau. Semicircle law on short scales and delocalization of eigenvectors for Wigner random matrices. *Annals of Probability*, 37:815–852, 2009.
- [5] Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, first edition, 1985.
- [6] L. V. Kantorovich. Functional analysis and applied mathematics. *Uspekhi Matematicheskikh Nauk*, 3(6):89–185, 1948. English translation available in L. V. Kantorovich, *Selected Works*, vol. 2, 171–280, (1996), Gordon and Breach Science Publishers.
- [7] V. Kargin. A concentration inequality and a local law for the sum of two random matrices. *Probability Theory and Related Fields*, 2011. Accepted for publication.
- [8] V. Kargin. An inequality for the distance between densities of free convolutions. available at <http://arxiv.org/abs/1107.0477>, 2011.
- [9] J. E. Littlewood. On inequalities in the theory of functions. *Proceedings of the London Mathematical Society*, 23:481–519, 1925.

- [10] L. Pastur and V. Vasilchuk. On the law of addition of random matrices. *Communications in Mathematical Physics*, 214:249–286, 2000.
- [11] Roland Speicher. Free convolution and the random sum of matrices. *Publications of RIMS (Kyoto University)*, 29:731–744, 1993.
- [12] Dan Voiculescu. Limit laws for random matrices and free products. *Inventiones mathematicae*, 104:201–220, 1991.
- [13] Dan Voiculescu. The analogues of entropy and of Fisher/s information measure in free probability theory I. *Communications in Mathematical Physics*, 155:71–92, 1993.